

CRF Errors Edited by the STIC Systems  
Branch

Serial Number: 10/505,405

CRF Edit Date: 9/2/04  
Edited by: Ro

Realigned nucleic acid/amino acid numbers/text in cases where the sequence  
text "wrapped" to the next line

Corrected the SEQ ID NO. Sequence numbers edited were:

ENTERED

Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID  
NO's edited:

☒ Deleted: ☒ invalid beginning/end-of-file text ; ☐ page numbers

Inserted mandatory headings/numeric identifiers, specifically:

Moved responses to same line as heading/numeric identifier, specifically:

Other:



PCT

## RAW SEQUENCE LISTING

DATE: 09/02/2004

PATENT APPLICATION: US/10/505,405

TIME: 12:24:07

Input Set : A:\pto.pg.txt

Output Set: N:\CRF4\09012004\J505405.raw

```

3 <110> APPLICANT: COLE, Stewart
4     BROSCHE, Roland
5     GORDON, Stephen
6     EIGLMEIER, Karin
7     GARNIER, Thierry
9 <120> TITLE OF INVENTION: DELETED SEQUENCE IN M. TUBERCULOSIS, METHOD FOR
10    DETECTING MYCOBACTERIA USING THESE SEQUENCES AND
11    VACCINES
13 <130> FILE REFERENCE: 346 379- US
C--> 15 <140> CURRENT APPLICATION NUMBER: US/10/505,405
C--> 15 <141> CURRENT FILING DATE: 2004-08-24
15 <150> PRIOR APPLICATION NUMBER: PCT/IB 03/00 986
16 <151> PRIOR FILING DATE: 2003-02-25
18 <150> PRIOR APPLICATION NUMBER: EP 02/290 458
19 <151> PRIOR FILING DATE: 2002-02-25
21 <160> NUMBER OF SEQ ID NOS: 22
23 <170> SOFTWARE: PatentIn Ver. 2.1
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 3953
27 <212> TYPE: DNA
28 <213> ORGANISM: Mycobacterium tuberculosis strain 74 ("ancestral" strain)
30 <220> FEATURE:
31 <221> NAME/KEY: CDS
32 <222> LOCATION: (735)..(3638)
34 <400> SEQUENCE: 1
35 tccagcgcg ccatcagcga tgaactctgg gacctgctac ccggctacct catcttccgg 60
37 tccatcatcc ccaaccggcc gccacccag gacacggtgc aagccctcgt cgacgacgtg 120
39 atactcccca gcctcaccgg atccaccggt tgagtcagcg gtgcgaatgg ctgggcaccg 180
41 ttgtggtgtc cgggtcccga ccgtactgtt gaatccgagg atccccgcct gaggtacggg 240
43 gcgtggtcgc gccccgggca atagcgtcgc cggttatcga aaggctaacg ggtgcagggg 300
45 atttcagtga ctggcctggt caaacgcggc tggatggtgc tggttgccgt ggcggtggtg 360
47 gcggtcgcgg gattcagcgt ctatcggttg cacggcatct tcggctcgca cgacaccacc 420
49 tcgaccgccg gtggtgtcgc gaacgacatc aagccgttca accccaaaca ggtaaccctc 480
51 gaggtctttg gcgctcccgg aaccgtggca acgatcaatt atctggacgt ggatgccaca 540
53 cctcggcaag tcctggacac gaccctgccc tggtcataca cgatcacgac gaccctgccc 600
55 gcggtcttcg ccaatgttgt cgcgcaaggc gacagcaatt ccatcggtcg ccgcatcacc 660
57 gtcaacggtg tagtcaagga cgaaaggatc gtcaacgaag tgcgcgccta taccttctgc 720
59 ctcgacaagt cctc atg agc aac cac cac cgc ccg ccg cct tgg ttg ccg 770
60      Met Ser Asn His His Arg Pro Arg Pro Trp Leu Pro
61      1              5              10
63 cac acc atc cga cgg ctt tcg ttg ccg atc ttg ctg ttt tgg gtg ggt 818
64 His Thr Ile Arg Arg Leu Ser Leu Pro Ile Leu Leu Phe Trp Val Gly
65      15              20              25

```

## RAW SEQUENCE LISTING

DATE: 09/02/2004

PATENT APPLICATION: US/10/505,405

TIME: 12:24:07

Input Set : A:\pto.pg.txt

Output Set: N:\CRF4\09012004\J505405.raw

```

67 gtc gcc gcc ata acc aat gcc gcc gtc ccg caa ttg gag gtc gtc ggg 866
68 Val Ala Ala Ile Thr Asn Ala Ala Val Pro Gln Leu Glu Val Val Gly
69      30      35      40
71 gag gcg cat aac gtc gca cag agc tcc ccg gat gac ccg tcg ctg cag 914
72 Glu Ala His Asn Val Ala Gln Ser Ser Pro Asp Asp Pro Ser Leu Gln
73 45      50      55      60
75 gcg atg aaa cgc atc ggc aag gtc ttc cac gag ttc gat tcc gac agt 962
76 Ala Met Lys Arg Ile Gly Lys Val Phe His Glu Phe Asp Ser Asp Ser
77      65      70      75
79 gcg gcc atg atc gtc ttg gaa ggc gat aag ccg ctc ggc aac gac gcc 1010
80 Ala Ala Met Ile Val Leu Glu Gly Asp Lys Pro Leu Gly Asn Asp Ala
81      80      85      90
83 cac ccg ttc tac gac acc ctg ctc cgc aac ctt tca aac gac acc aaa 1058
84 His Arg Phe Tyr Asp Thr Leu Leu Arg Asn Leu Ser Asn Asp Thr Lys
85      95      100      105
87 cac gtc gag cac gtt cag gac ttc tgg ggc gat ccg ctg acc gcg gcc 1106
88 His Val Glu His Val Gln Asp Phe Trp Gly Asp Pro Leu Thr Ala Ala
89      110      115      120
91 ggc tcg caa agc acc gac ggc aaa gcc gcc tac gtt cag gtc tat ctc 1154
92 Gly Ser Gln Ser Thr Asp Gly Lys Ala Ala Tyr Val Gln Val Tyr Leu
93 125      130      135      140
95 gcc gcc aac caa ggc gag gcg ttg tca atc gag tcc gtc gac gcg gtc 1202
96 Ala Gly Asn Gln Gly Glu Ala Leu Ser Ile Glu Ser Val Asp Ala Val
97      145      150      155
99 ccg gac atc gtc gcc cat acg cca cca ccg gcc ggc gtc aag gcc tac 1250
100 Arg Asp Ile Val Ala His Thr Pro Pro Pro Ala Gly Val Lys Ala Tyr
101      160      165      170
103 gtc acc ggc gcg gcc ccg ctc atg gcc gat cag ttt cag gtc ggc agc 1298
104 Val Thr Gly Ala Ala Pro Leu Met Ala Asp Gln Phe Gln Val Gly Ser
105      175      180      185
107 aaa gga acc gcg aaa gtt acc ggg ata act ctg gtt gtc atc gcg gtc 1346
108 Lys Gly Thr Ala Lys Val Thr Gly Ile Thr Leu Val Val Ile Ala Val
109      190      195      200
111 atg ttg ctc ttc gta tac cgt tcc gtc gtc acc atg gtc ctg gtc ctt 1394
112 Met Leu Leu Phe Val Tyr Arg Ser Val Val Thr Met Val Leu Val Leu
113 205      210      215      220
115 atc acg gtt ctt att gag ttg gcc gcg gcc ccg ggc atc gtc gct ttt 1442
116 Ile Thr Val Leu Ile Glu Leu Ala Ala Arg Gly Ile Val Ala Phe
117      225      230      235
119 ctc gga aac gcc ggc gta atc ggc ctg tcg aca tac tcg acg aat ctg 1490
120 Leu Gly Asn Ala Gly Val Ile Gly Leu Ser Thr Tyr Ser Thr Asn Leu
121      240      245      250
123 ctc aca cta ttg gta atc gcg gcg ggc aca gac tac gcg att ttt gtc 1538
124 Leu Thr Leu Leu Val Ile Ala Ala Gly Thr Asp Tyr Ala Ile Phe Val
125      255      260      265
127 ctc ggc cgc tat cac gag gcg cgc tac gcc gca cag gat ccg gaa acg 1586
128 Leu Gly Arg Tyr His Glu Ala Arg Tyr Ala Ala Gln Asp Arg Glu Thr
129      270      275      280
131 gcc ttc tac acg atg tat cgc ggc acc gcc cac gtc gtc ttg ggc tcg 1634

```

## RAW SEQUENCE LISTING

DATE: 09/02/2004

PATENT APPLICATION: US/10/505,405

TIME: 12:24:07

Input Set : A:\pto.pg.txt

Output Set: N:\CRF4\09012004\J505405.raw

132	Ala	Phe	Tyr	Thr	Met	Tyr	Arg	Gly	Thr	Ala	His	Val	Val	Leu	Gly	Ser	
133	285					290					295					300	
135	ggt	ctg	acc	gtt	gcc	ggc	gcg	gtg	tat	tgc	ctg	agc	ttt	acc	cgg	cta	1682
136	Gly	Leu	Thr	Val	Ala	Gly	Ala	Val	Tyr	Cys	Leu	Ser	Phe	Thr	Arg	Leu	
137					305					310					315		
139	ccc	tat	ttt	caa	agc	ctg	ggt	att	ccc	gcc	tcg	ata	ggg	gtg	atg	att	1730
140	Pro	Tyr	Phe	Gln	Ser	Leu	Gly	Ile	Pro	Ala	Ser	Ile	Gly	Val	Met	Ile	
141				320					325					330			
143	gcg	ttg	gca	gcc	gcg	ctc	agc	ctg	gcc	cca	tcc	gtg	ctc	atc	ttg	ggc	1778
144	Ala	Leu	Ala	Ala	Ala	Leu	Ser	Leu	Ala	Pro	Ser	Val	Leu	Ile	Leu	Gly	
145			335						340				345				
147	agt	cgt	ttc	ggt	tgt	ttc	gaa	ccc	aag	cgc	agg	atg	agg	acc	agg	gga	1826
148	Ser	Arg	Phe	Gly	Cys	Phe	Glu	Pro	Lys	Arg	Arg	Met	Arg	Thr	Arg	Gly	
149		350					355				360						
151	tgg	cgg	cgc	atc	ggc	acg	gcc	atc	gtg	cgt	tgg	ccg	gga	ccc	atc	ctg	1874
152	Trp	Arg	Arg	Ile	Gly	Thr	Ala	Ile	Val	Arg	Trp	Pro	Gly	Pro	Ile	Leu	
153	365				370					375					380		
155	gca	gtg	gcg	tgc	gca	att	gcg	gtg	gtg	ggt	ctg	ctc	gcg	ctg	ccg	gga	1922
156	Ala	Val	Ala	Cys	Ala	Ile	Ala	Val	Val	Gly	Leu	Leu	Ala	Leu	Pro	Gly	
157				385						390					395		
159	tac	aaa	acg	agc	tac	gac	gct	cgc	tat	tac	atg	ccc	gcc	acc	gcc	ccg	1970
160	Tyr	Lys	Thr	Ser	Tyr	Asp	Ala	Arg	Tyr	Tyr	Met	Pro	Ala	Thr	Ala	Pro	
161				400					405					410			
163	gcc	aat	att	ggc	tac	atg	gcc	gcg	gag	cga	cat	ttt	ccc	caa	gcg	cgg	2018
164	Ala	Asn	Ile	Gly	Tyr	Met	Ala	Ala	Glu	Arg	His	Phe	Pro	Gln	Ala	Arg	
165			415					420					425				
167	ctg	aat	ccc	gaa	cta	ctg	atg	atc	gag	acg	gat	cac	gat	atg	cgc	aat	2066
168	Leu	Asn	Pro	Glu	Leu	Leu	Met	Ile	Glu	Thr	Asp	His	Asp	Met	Arg	Asn	
169		430					435				440						
171	ccg	gcc	gac	atg	ctc	atc	ttg	gat	agg	atc	gcc	aag	gct	gtc	ttc	cat	2114
172	Pro	Ala	Asp	Met	Leu	Ile	Leu	Asp	Arg	Ile	Ala	Lys	Ala	Val	Phe	His	
173	445				450					455					460		
175	ctg	ccc	ggc	ata	ggg	ctg	gtg	cag	gcc	atg	acc	cgg	ccg	cta	gga	acc	2162
176	Leu	Pro	Gly	Ile	Gly	Leu	Val	Gln	Ala	Met	Thr	Arg	Pro	Leu	Gly	Thr	
177				465					470						475		
179	ccg	att	gac	cac	agc	tcg	ata	ccg	ttt	cag	atc	agc	atg	caa	agc	gtc	2210
180	Pro	Ile	Asp	His	Ser	Ser	Ile	Pro	Phe	Gln	Ile	Ser	Met	Gln	Ser	Val	
181				480					485					490			
183	ggc	cag	att	cag	aat	ctc	aag	tat	cag	agg	gac	cga	gca	gcc	gac	ttg	2258
184	Gly	Gln	Ile	Gln	Asn	Leu	Lys	Tyr	Gln	Arg	Asp	Arg	Ala	Ala	Asp	Leu	
185			495					500					505				
187	ctg	aag	cag	gcc	gaa	gag	ctg	ggg	aag	acg	atc	gaa	atc	ttg	cag	cgc	2306
188	Leu	Lys	Gln	Ala	Glu	Glu	Leu	Gly	Lys	Thr	Ile	Glu	Ile	Leu	Gln	Arg	
189		510					515				520						
191	caa	tat	gcc	cta	cag	cag	gaa	ctc	gcg	gcc	gct	act	cac	gag	caa	gcc	2354
192	Gln	Tyr	Ala	Leu	Gln	Gln	Glu	Leu	Ala	Ala	Ala	Thr	His	Glu	Gln	Ala	
193	525				530					535					540		
195	gaa	agc	ttt	cac	caa	acg	atc	gcc	acg	gta	aac	gaa	ctg	cga	gat	agg	2402
196	Glu	Ser	Phe	His	Gln	Thr	Ile	Ala	Thr	Val	Asn	Glu	Leu	Arg	Asp	Arg	

## RAW SEQUENCE LISTING

DATE: 09/02/2004

PATENT APPLICATION: US/10/505,405

TIME: 12:24:07

Input Set : A:\pto.pg.txt

Output Set: N:\CRF4\09012004\J505405.raw

197				545				550				555					
199	atc	gcc	aat	ttc	gac	gat	ttc	ttc	agg	ccg	att	cgt	agt	tac	ttt	tac	2450
200	Ile	Ala	Asn	Phe	Asp	Asp	Phe	Phe	Arg	Pro	Ile	Arg	Ser	Tyr	Phe	Tyr	
201				560				565				570					
203	tgg	gaa	aag	cac	tgc	tac	gat	atc	ccg	agc	tgc	tgg	gcg	ctg	aga	tcc	2498
204	Trp	Glu	Lys	His	Cys	Tyr	Asp	Ile	Pro	Ser	Cys	Trp	Ala	Leu	Arg	Ser	
205			575				580					585					
207	gtc	ttt	gac	acg	atc	gac	ggg	atc	gac	caa	ctc	ggc	gag	cag	ctg	gcc	2546
208	Val	Phe	Asp	Thr	Ile	Asp	Gly	Ile	Asp	Gln	Leu	Gly	Glu	Gln	Leu	Ala	
209		590					595					600					
211	agc	gtg	acc	gta	acc	ttg	gac	aag	ttg	gct	gcg	atc	cag	cct	caa	ttg	2594
212	Ser	Val	Thr	Val	Thr	Leu	Asp	Lys	Leu	Ala	Ala	Ile	Gln	Pro	Gln	Leu	
213	605					610				615					620		
215	gtg	gcg	ctg	cta	cca	gac	gag	atc	gcc	agc	cag	cag	atc	aat	cgg	gaa	2642
216	Val	Ala	Leu	Leu	Pro	Asp	Glu	Ile	Ala	Ser	Gln	Gln	Ile	Asn	Arg	Glu	
217				625				630				635					
219	ctg	gcg	ctg	gct	aac	tac	gcc	acc	atg	tcc	ggg	atc	tat	gcc	cag	acg	2690
220	Leu	Ala	Leu	Ala	Asn	Tyr	Ala	Thr	Met	Ser	Gly	Ile	Tyr	Ala	Gln	Thr	
221			640				645					650					
223	gcg	gcc	ttg	atc	gaa	aac	gct	gcc	gcc	atg	gga	caa	gcc	ttt	gac	gcc	2738
224	Ala	Ala	Leu	Ile	Glu	Asn	Ala	Ala	Ala	Met	Gly	Gln	Ala	Phe	Asp	Ala	
225			655				660					665					
227	gcc	aag	aac	gac	gac	tcc	ttc	tat	ctg	ccg	ccg	gag	gct	ttt	gac	aac	2786
228	Ala	Lys	Asn	Asp	Asp	Ser	Phe	Tyr	Leu	Pro	Pro	Glu	Ala	Phe	Asp	Asn	
229		670				675					680						
231	cca	gat	ttc	cag	cg	ggc	ctg	aaa	ttg	ttc	ctg	tcg	gca	gac	ggg	aag	2834
232	Pro	Asp	Phe	Gln	Arg	Gly	Leu	Lys	Leu	Phe	Leu	Ser	Ala	Asp	Gly	Lys	
233	685					690				695					700		
235	gcg	gct	cgg	atg	atc	atc	tcc	cat	gaa	ggc	gat	ccc	gcc	acc	ccc	gaa	2882
236	Ala	Ala	Arg	Met	Ile	Ile	Ser	His	Glu	Gly	Asp	Pro	Ala	Thr	Pro	Glu	
237			705				710					715					
239	ggc	att	tcg	cat	atc	gac	gcg	atc	aag	cag	gcg	gcc	cac	gag	gcc	gtg	2930
240	Gly	Ile	Ser	His	Ile	Asp	Ala	Ile	Lys	Gln	Ala	Ala	His	Glu	Ala	Val	
241			720				725					730					
243	aag	ggc	act	ccc	atg	gcg	ggg	gct	ggg	atc	tat	ctg	gcc	ggc	acg	gcc	2978
244	Lys	Gly	Thr	Pro	Met	Ala	Gly	Ala	Gly	Ile	Tyr	Leu	Ala	Gly	Thr	Ala	
245			735				740					745					
247	gcc	acc	ttc	aag	gac	att	caa	gac	ggc	gcc	acc	tac	gac	ctc	ctg	atc	3026
248	Ala	Thr	Phe	Lys	Asp	Ile	Gln	Asp	Gly	Ala	Thr	Tyr	Asp	Leu	Leu	Ile	
249		750				755					760						
251	gcc	gga	ata	gcc	gcg	ctg	agc	ttg	att	ttg	ctc	atc	atg	atg	atc	att	3074
252	Ala	Gly	Ile	Ala	Ala	Leu	Ser	Leu	Ile	Leu	Leu	Ile	Met	Met	Ile	Ile	
253	765					770				775					780		
255	acc	cga	agc	ctg	gtt	gcg	gcg	ctg	gtg	atc	gtg	ggc	acg	gtg	gcg	ctg	3122
256	Thr	Arg	Ser	Leu	Val	Ala	Ala	Leu	Val	Ile	Val	Gly	Thr	Val	Ala	Leu	
257			785				790					795					
259	tcg	ttg	ggc	gct	tct	ttt	ggc	ctg	tcc	gtg	ctg	gtg	tgg	cag	cat	ctt	3170
260	Ser	Leu	Gly	Ala	Ser	Phe	Gly	Leu	Ser	Val	Leu	Val	Trp	Gln	His	Leu	
261			800				805					810					

## RAW SEQUENCE LISTING

DATE: 09/02/2004

PATENT APPLICATION: US/10/505,405

TIME: 12:24:07

Input Set : A:\pto.pg.txt

Output Set: N:\CRF4\09012004\J505405.raw

```

263 ctc ggt atc cag ttg tac tgg atc gtg ctc gcg ctg gcc gtc atc ctg 3218
264 Leu Gly Ile Gln Leu Tyr Trp Ile Val Leu Ala Leu Ala Val Ile Leu
265      815      820      825
267 ctc ctg gcc gtg gga tcg gac tat aac ttg ctg ctg att tcc cga ttc 3266
268 Leu Leu Ala Val Gly Ser Asp Tyr Asn Leu Leu Leu Ile Ser Arg Phe
269      830      835      840
271 aag gag gag atc ggt gca ggt ttg aac acc ggc atc atc cgt gcg atg 3314
272 Lys Glu Glu Ile Gly Ala Gly Leu Asn Thr Gly Ile Ile Arg Ala Met
273 845      850      855      860
275 gcc ggc acc ggc ggg gtg gtg acc gct gcc ggc ctg gtg ttc gcc gcc 3362
276 Ala Gly Thr Gly Gly Val Val Thr Ala Ala Gly Leu Val Phe Ala Ala
277      865      870      875
279 act atg tct tcg ttc gtg ttc agt gat ttg cgg gtc ctc ggt cag atc 3410
280 Thr Met Ser Ser Phe Val Phe Ser Asp Leu Arg Val Leu Gly Gln Ile
281      880      885      890
283 ggg acc acc att ggt ctt ggg ctg ctg ttc gac acg ctg gtg gtg cgc 3458
284 Gly Thr Thr Ile Gly Leu Gly Leu Leu Phe Asp Thr Leu Val Val Arg
285      895      900      905
287 gcg ttc atg acc ccg tcc atc gcg gtg ctg ctc ggg cgc tgg ttc tgg 3506
288 Ala Phe Met Thr Pro Ser Ile Ala Val Leu Leu Gly Arg Trp Phe Trp
289      910      915      920
291 tgg ccg caa cga gtg cgc ccg cgc cct gcc agc agg atg ctt cgg ccg 3554
292 Trp Pro Gln Arg Val Arg Pro Arg Pro Ala Ser Arg Met Leu Arg Pro
293 925      930      935      940
295 tac ggc ccg cgg ccc gtg gtt cgt gaa ttg ctg ctg cgc gag ggc aac 3602
296 Tyr Gly Pro Arg Pro Val Val Arg Glu Leu Leu Leu Arg Glu Gly Asn
297      945      950      955
299 gat gac ccg aga act cag gtg gct acc cac cgt taa ggtggtggga 3648
300 Asp Asp Pro Arg Thr Gln Val Ala Thr His Arg
301      960      965
303 tgccgctttc aggggaatat gcgccgagcc cgctcgactg gtcgcgcgag caagccgaca 3708
305 cgtatatgaa gtccggcgga accgagggca cacagctgca gggaaagccg gtcctcctgc 3768
307 tcaccaccgt cggggcgaag accggcaaac tccgtaagac cccgctgatg cgcgtcgagc 3828
309 acgacggcca gtacgcgatc gtcgcctcgc tgggtggggc gccgaaaaat ccggtctggt 3888
311 accacaacgt cgtgaagaac ccacgggtcg agctgcagga cggcaccgga ccggcgacta 3948
313 cgacg 3953
316 <210> SEQ ID NO: 2
317 <211> LENGTH: 967
318 <212> TYPE: PRT
319 <213> ORGANISM: Mycobacterium tuberculosis strain 74 ("ancestral" strain)
321 <220> FEATURE:
322 <223> OTHER INFORMATION: mmpL6 protein
324 <400> SEQUENCE: 2
325 Met Ser Asn His His Arg Pro Arg Pro Trp Leu Pro His Thr Ile Arg
326 1 5 10 15
328 Arg Leu Ser Leu Pro Ile Leu Leu Phe Trp Val Gly Val Ala Ala Ile
329 20 25 30
331 Thr Asn Ala Ala Val Pro Gln Leu Glu Val Val Gly Glu Ala His Asn
332 35 40 45

```

## VERIFICATION SUMMARY

DATE: 09/02/2004

PATENT APPLICATION: US/10/505,405

TIME: 12:24:08

Input Set : A:\pto.pg.txt

Output Set: N:\CRF4\09012004\J505405.raw

L:15 M:270 C: Current Application Number differs, Replaced Current Application No

L:15 M:271 C: Current Filing Date differs, Replaced Current Filing Date